Asymptotic Optimality of Approximate Filters in Stochastic Systems with Colored Noises

A. Le Breton, M. C. Roubaud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Asymptotic Optimality of Approximate Filters in Stochastic Systems with Colored Noises

A. Le Breton and M. C. Roubaud

N° 3801

Novembre 1999

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--3801--FR+ENG

Asymptotic Optimality of Approximate Filters in Stochastic Systems with Colored Noises

A. Le Breton* and M. C. Roubaud†

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet SYSDYS

Rapport de recherche n° 3801 — Novembre 1999 — 35 pages

Abstract: In this report, a new interpretation of the main part of our study on asymptotic behaviour of reduced-order filters in [6] is given. Approximate filters are proposed for semi-linear and nonlinear stochastic systems with colored noises. Basically these filters are defined as those which are optimal when the noises are white. Their long time behavior is investigated and their asymptotic optimality is shown in two cases under some reasonable assumptions. At first the case of a system where the signal and observation dynamics are linear with respect to the state is considered; the approximate filter is a Kalman filter and the asymptotic analysis relies on a representation of the optimal filter which extends a formula known for a linear system with white noises and non-Gaussian initial condition. Then the case of a nonlinear system with limiting ergodic behavior is analyzed; here the approximate filter appears as the optimal filter corresponding to an incorrect prior distribution and the asymptotic study uses the results of Ocone and Pardoux on the asymptotic stability of optimal filters with respect to their initial condition.

Key-words: nonlinear filtering, colored noises, approximate filters, asymptotic optimality

* LMC–UNIVERSITÉ J. FOURIER, BP 53, 38041 Grenoble Cedex 9 (France), `Alain.Le-Breton@imag.fr`
† SYSDYS–LATP, 39 rue Joliot–Curie, 13453 Marseille Cedex 13 (France), `mcroubau@sophia.inria.fr`

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Filtres approchés asympotiquement optimaux pour des systémes avec bruit colorés

Résumé : Dans ce rapport nous donnons une nouvelle interprétation de la majeure partie de notre étude sur le comportement asymptotique de filtres d'ordre réduit présentée dans [6]. Nous proposons des filtres approchés pour des systèmes stochastiques semi-linéaires et non linéaires avec des bruits colorés. Ces filtres sont définis comme étant identiques aux filtres optimaux lorsque les bruits sont blancs. Nous étudions leur comportement en temps long et nous montrons leur efficacité asymptotique dans deux situations sous des hypothèses raisonables. Premièrement, nous considérons le cas d'un système où la dynamique du signal et celle de l'observation sont linéaires; le filtre approché est un filtre de Kalman et dans l'étude asymptotique nous donnons une représentation du filtre optimal. Ceci étend une formule connue pour des systèmes avec bruits blancs et condition initiale non gaussienne. Puis nous considérons le cas d'un système non linéaire sous des hypothèses d'ergodicité. Dans cette situation le filtre approché correspond au filtre optimal du système initialisé avec une distribution erronée. Pour l'étude asymptotique nous utilisons les résultats d'Ocone et Pardoux sur la stabilité asymptotique du filtre optimal par rapport à la condition initiale.

Mots-clés : Filtrage non linéaire, bruits colorés, filtres approchés, efficacité asymptotique

Contents

1 Introduction.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space. Let $X = (X_t,\, t \geq 0)$ and $Y = (Y_t,\, t \geq 0)$ be processes, taking values in $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively, governed by the system of stochastic differential equations

$$\begin{cases} dX_t &= b(X_t)dt + d\xi_t,\ t \geq 0,\ X_0, \\ dY_t &= h(X_t)dt + d\eta_t,\ t \geq 0,\ Y_0 = 0. \end{cases} \tag{1}$$

Here b and h are functions from $\mathbb{R}^d$ into $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively. Moreover X_0 stands for some $\mathcal{F}_0$-measurable random initial condition for X, with a given distribution ν_0, and $\xi = (\xi_t,\, t \geq 0)$ and $\eta = (\eta_t,\, t \geq 0)$ are $(\mathcal{F}_t)$-adapted noise processes. Let us assume that the system (1) has a uniquely defined solution. Supposing that only Y is observed but one wishes to know X, the classical problem of filtering the signal X at time t from the observation of Y up to time t occurs. The solution to this problem is the conditional distribution Π_t of X_t given $\mathcal{Y}_t = \sigma(\{Y_s\,,\, 0 \leq s \leq t\})$ which we shall call the *exact filter.*

It is well known that if in model (1) the functions b and h are linear, the processes ξ and η are $(\mathcal{F}_t)$-Brownian motions and the distribution ν_0 is Gaussian with mean m_0 and covariance Q_0, then the exact filter is a *Kalman filter.* More precisely it is a Gaussian distribution whose mean and covariance are the unique solutions to Kalman filtering equations initialized with m_0 and Q_0. Moreover when the initial condition X_0 in (1) is non-Gaussian, then the exact filter can be expressed in terms of some family of Kalman filters (see e.g. Makowski [7] and Beněs, Karatzas [1]). Actually Ocone and Pardoux [8] showed that, under stabilizability and detectability assumptions, whatever is the distribution of X_0 provided that $\mathbb{E}[|X_0|^2] < +\infty$, asymptotically in time the exact filter is approached by a Kalman filter, i.e. a Gaussian distribution whose mean and covariance are governed by the Kalman filtering equations initialized with any vector in $\mathbb{R}^d$ and any symmetric nonnegative definite $d \times d$ matrix. In the same paper Ocone and Pardoux proved also that in the case of a signal with limiting ergodic behavior observed in additive white noise through a nonlinear channel, the solution to the filtering equations initialized with an incorrect prior distribution approaches the exact filter as time goes to infinity.

In the present paper we consider the filtering model (1) when the processes ξ and η are possibly non-independent and non-Gaussian colored noises.

We assume that

$$\begin{cases} d\xi_t &= b_\zeta(\zeta_t)dt + dV_t, \ t \geq 0, \ \xi_0 = 0, \\ d\eta_t &= h_\zeta(\zeta_t)dt + dW_t, \ t \geq 0, \ \eta_0 = 0. \end{cases} \tag{2}$$

Here $V = (V_t, t \geq 0)$ and $W = (W_t, t \geq 0)$ are independent $(\mathcal{F}_t)$-Brownian motions in $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively. We assume that $\mathbb{E}[V_1 V_1'] = \Sigma_V$ (resp. $\mathbb{E}[W_1 W_1'] = \Sigma_W$) for some given symmetric nonnegative (resp. positive) definite matrix Σ_V (resp. Σ_W, which actually without loss of generality will be taken as the $p \times p$ identity matrix I_p). Moreover $\zeta = (\zeta_t, t \geq 0)$ is a $(\mathcal{F}_t)$-adapted perturbation process taking values in $\mathbb{R}^k$ and of course b_ζ and h_ζ are functions from $\mathbb{R}^k$ into $\mathbb{R}^d$ and $\mathbb{R}^p$ respectively. We suppose also that the pair (ζ, X_0) is independent of (V, W). Of course in model (1)-(2) the exact filter Π_t for X is nothing but the marginal distribution of the exact filter for the pair (ζ, X). The determination of Π_t, which is in general an infinite dimensional problem, requires the complete knowledge of the joint probabilistic structure of ζ, X and Y and leads to computations in a space dimension which is larger than that of the system of interest. Hence it is interesting to elaborate *approximate filters* which ignore a part of the actual structure of the system, and then are more easily computable, but may have a long time behavior close to that of the exact filter. Here this question is addressed under the assumptions that

- (A_0) $\quad b_\zeta(0) = 0, \quad h_\zeta(0) = 0,$

and the perturbation process ζ converges to zero in some appropriate sense which will be made precise later. We propose approximate filters which are basically the optimal filters when the noises are independent white noises, i.e. $\zeta \equiv 0$ since then, from (A_0), $\xi \equiv V$ and $\eta \equiv W$. The asymptotic optimality of these approximate filters is shown in some cases under reasonable additional assumptions.

The paper is organized as follows. At first in Section 2 some notations are fixed, preliminaries concerning the model and filters are discussed and a particular case where all functions b, b_ζ, h and h_ζ in model (1)-(2) are linear is analyzed. Then in Section 3 we consider what we call the *semi-linear case* where the functions b and h are still linear but functions b_ζ and h_ζ may be nonlinear. We derive a representation of the exact filter which extends the formula obtained by Makowski [7] and Beněs and Karatzas [1] for a linear system with white noises and non-Gaussian initial conditions. Here the approximate filter can be taken as a Kalman filter and the asymptotic analysis is led under some stabilizability and detectability assumptions. Finally in Section 4 a *nonlinear case* where the signal has a limiting ergodic behavior is investigated. Here the approximate filter appears as the optimal filter corresponding to an incorrect prior distribution and the asymptotic study uses the results of Ocone and Pardoux [8] on the asymptotic stability of optimal filters with respect to their initial condition.

2 Preliminaries.

Throughout, concerning the filtering model (1)-(2), we assume the following: either

- (A_1) $\quad b, b_\zeta, h$ and h_ζ are Lipschitzian functions,

or

- (A'_1) $\quad b, b_\zeta$ and h_ζ are Lipschitzian; h is continuous; h_ζ and h are bounded.

We suppose also that the paths of the process ζ are almost surely locally square integrable i.e.

- (A_2) $\quad$ for any $t > 0$ $\quad \mathbb{P}([\int_0^t |\zeta_s|^2\, ds < +\infty]) = 1.$

Then we can assert that the solution process (X, Y) of the system described by (1)-(2) is uniquely defined on the given stochastic basis.

2.1 Filters.

Actually it is convenient to work on the canonical probability space of the process (ζ, X, Y). Let us define $\Omega_0 = \mathcal{L}^2_{loc}(\mathbb{R}^+; \mathbb{R}^k)$ the set of locally square integrable functions from $\mathbb{R}^+$ into $\mathbb{R}^k$ and $\Omega_1 = \mathcal{C}(\mathbb{R}^+ ; \mathbb{R}^d)$ (resp. $\Omega_2 = \mathcal{C}_0(\mathbb{R}^+ ; \mathbb{R}^p)$) the set of continuous functions from $\mathbb{R}^+$ into $\mathbb{R}^d$ (resp. $\mathbb{R}^p$ with value zero at $t = 0$). Define $\Omega = \Omega_0 \times \Omega_1 \times \Omega_2$, $\mathcal{F}$ its Borel field, $(\mathcal{F}_t)$ the canonical Borel filtration on Ω and $(\zeta_t, X_t, Y_t)(z, x, y) = (z_t, x_t, y_t)$ for any $(z, x, y) \in \Omega$. Let $\mathbb{P}$ be the distribution on Ω corresponding to the filtering model defined by (1)-(2). Then in particular, on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, the processes V and W defined by

$$\left\{ \begin{array}{lcl} V_t & \triangleq & X_t - X_0 - \int_0^t b(X_s)ds - \int_0^t b_\zeta(\zeta_s)ds, \;\; t \geq 0, \\ W_t & \triangleq & Y_t - \int_0^t h(X_s)ds - \int_0^t h_\zeta(\zeta_s)ds, \;\; t \geq 0, \end{array} \right. \tag{3}$$

are independent Brownian motions in $\mathbb{R}^d$ and in $\mathbb{R}^p$ such that $\mathbb{E}[V_1 V_1'] = \Sigma_V$ and $\mathbb{E}[W_1 W_1'] = I_p$ respectively. Moreover (ζ, X_0) is independent of (V, W). Similarly let $\mathbb{P}^0$ be the distribution on Ω corresponding to the filtering model defined by (1)-(2) when the noises are white i.e. $\zeta \equiv 0$ since we assume that (A_0) is fulfilled. Of course $\mathbb{P}^0$ is nothing but the product probability $\delta_0 \otimes P$ where δ_0 is the Dirac measure on Ω_0 concentrated at 0 and P is the distribution on $\Omega_1 \times \Omega_2$ corresponding to the filtering model (1) when ξ and η are independent Brownian motions and the distribution of X_0 is ν_0. In what follows the symbols $\mathbb{E}[\,.\,]$ and $\mathbb{E}^0[\,.\,]$ will stand for an expectation or a conditional expectation computed with respect to $\mathbb{P}$ and $\mathbb{P}^0$ respectively. So for the exact filter Π_t of X_t, i.e. the conditional distribution of X_t given $\mathcal{Y}_t$ with respect to $\mathbb{P}$, we write

$$\Pi_t(\varphi) \triangleq \mathbb{E}[\varphi(X_t) \,|\, \mathcal{Y}_t],$$

for any φ in the set $\mathcal{M}_b(\mathbb{R}^d)$ of all bounded measurable functions from $\mathbb{R}^d$ into $\mathbb{R}$ or in the set $\mathcal{M}^+(\mathbb{R}^d)$ of all nonnegative measurable functions from $\mathbb{R}^d$ into $\mathbb{R}$. Similarly, when X_t is $\mathbb{P}$-integrable, for the conditional expectation we write

$$\hat{X}_t \triangleq \mathbb{E}[X_t \,|\, \mathcal{Y}_t].$$

According to the discussion in Section 1, the basic approximate filter is taken as the conditional distribution Π_t^0 of X_t given $\mathcal{Y}_t$ with respect to $\mathbb{P}^0$ and we write also

$$\Pi_t^0(\varphi) \triangleq \mathbb{E}^0[\varphi(X_t) \,|\, \mathcal{Y}_t],$$

for any function φ as above and, when it is well defined

$$\hat{X}_t^0 \triangleq \mathbb{E}^0[X_t \mid \mathcal{Y}_t].$$

Actually we shall also consider approximate filters defined similarly to Π_0 but with some more convenient prior distribution ν in place of the true one ν_0. This will be made precise later.

Now let us establish representations of filters which shall be used in Sections 3 and 4 below. Let us suppose here that the Brownian motion V is also standard.
From assumptions (A_1) or (A_1'), for any fixed $z = (z_t, t \geq 0) \in \Omega_0$ we may define the processes

$$K_t(z) \triangleq \int_0^t b_\zeta(z_s)'\, dV_s + \int_0^t h_\zeta(z_s)'\, dW_s,\ t \geq 0, \tag{4}$$

and

$$< K >_t (z) \triangleq \int_0^t |b_\zeta(z_s)|^2\, ds + \int_0^t |h_\zeta(z_s)|^2\, ds,\ t \geq 0. \tag{5}$$

Throughout for a vector $v \in \mathbb{R}^l$ (resp. a $l \times m$ matrix A), v' and $|v|$(resp. A' and $||A||$) stand for the transpose and the usual Euclidian norm. Due to assumption (A_3), for any $t > 0$ we may also introduce the random variables

$$K_t \triangleq K_t(\zeta); \quad < K >_t \triangleq < K >_t (\zeta); \quad L_t \triangleq \exp\{-K_t - \frac{1}{2} < K >_t\}. \tag{6}$$

Conditioning on ζ, it is easy to check that $\mathbb{E}[L_t] = 1$. Hence we may define the probability $\widetilde{\mathbb{P}}$ on Ω which is locally absolutely continuous with respect to $\mathbb{P}$ with the local Radon-Nikodym derivative L_t. Then, defining the processes $\widetilde{V} = (\widetilde{V}_t\,,\, t \geq 0)$ and $\widetilde{W} = (\widetilde{W}_t\,,\, t \geq 0)$ by

$$\begin{cases} \widetilde{V}_t & \triangleq & \int_0^t b_\zeta(\zeta_s)\, ds + V_t,\ t \geq 0, \\ \widetilde{W}_t & \triangleq & \int_0^t h_\zeta(\zeta_s)\, ds + W_t\,,\ t \geq 0, \end{cases} \tag{7}$$

applying the Girsanov Theorem, we get that with respect to the probability $\widetilde{\mathbb{P}}$ the process $\widetilde{V}$ and $\widetilde{W}$ are independent standard Brownian motions. Moreover the pair (ζ, X_0) is independent of $(\widetilde{V}, \widetilde{W})$ and the distribution of (ζ, X_0) under $\widetilde{\mathbb{P}}$ is the same as under $\mathbb{P}$, $\mathbb{P}_{(\zeta,X_0)}$ say. Then of course the marginal distribution of $\widetilde{\mathbb{P}}$ on $\Omega_1 \times \Omega_2$ is nothing but the probability P. Consequently, denoting by $\widetilde{\mathbb{E}}[\,.\,]$ an expectation computed with respect to $\widetilde{\mathbb{P}}$, the approximate filter Π_t^0 may be equivalently defined as

$$\Pi_t^0(\varphi) = \widetilde{\mathbb{E}}[\varphi(X_t) \,|\, \mathcal{Y}_t], \;\; \varphi \in \mathcal{M}^+(\mathbb{R}^d).$$

Concerning the exact filter, by the classical Bayes formula, we have

$$\Pi_t(\varphi) = \frac{\widetilde{\mathbb{E}}[\varphi(X_t) L_t^{-1} \,|\, \mathcal{Y}_t]}{\widetilde{\mathbb{E}}[L_t^{-1} \,|\, \mathcal{Y}_t]}, \;\; \varphi \in \mathcal{M}^+(\mathbb{R}^d). \tag{8}$$

The next lemma shows that actually if ζ and X_0 are independent then both the numerator and the denominator in the right hand side of (8) may be represented in terms of some family of filters with respect to $\widetilde{\mathbb{P}}$ (or equivalently $\mathbb{P}^0$). For any $z = (z_t, t \geq 0) \in \Omega_0$, we introduce the following filtering model indexed by z:

$$\left\{ \begin{array}{rcl} dX_t & = & b(X_t)dt + d\widetilde{V}_t;\; t \geq 0;\; X_0, \\ d\widetilde{K}_t(z) & = & b_\zeta(z_t)' d\widetilde{V}_t + h_\zeta(z_t)' d\widetilde{W}_t;\; t \geq 0;\; \widetilde{K}_0(z) = 0, \\ dY_t & = & h(X_t)dt + d\widetilde{W}_t;\; t \geq 0;\; Y = 0, \end{array} \right. \tag{9}$$

where $(X_t, \widetilde{K}_t(z))$ is the $(d+1)$-dimensional signal and Y is the observation process taking values in $\mathbb{R}^p$.

Now we can state

Lemma 2.1 *Assume that ζ and X_0 are independent. Then for any $\psi \in \mathcal{M}^+(\mathbb{R}^{k+d})$ the following equality holds*

$$\widetilde{\mathbb{E}}[\psi(\zeta_t, X_t) L_t^{-1} \,|\, \mathcal{Y}_t] = \int_{\Omega_0} d\mathbb{P}_\zeta(z) e^{-\frac{1}{2}<K>_t(z)} \left\{ \int_{\mathbb{R}^{d+1}} \psi(z_t, u) e^v \tilde{\pi}_t^z(du, dv) \right\}, \tag{10}$$

where $\mathbb{P}_\zeta$ denotes the law of ζ and, for all $z = (z_t, t \geq 0) \in \Omega_0$, the symbol $\tilde{\pi}_t^z$ stands for the conditional distribution of $(X_t, \widetilde{K}_t(z))$ given $\mathcal{Y}_t$ in model (9) taken under the probability $\widetilde{\mathbb{P}}$ (or $\mathbb{P}^0$).

Proof Due to the definition of L_t in (6), we get

$$\widetilde{\mathbb{E}}[\psi(\zeta_t, X_t)L_t^{-1} \,|\, \mathcal{Y}_t] = \widetilde{\mathbb{E}}[\psi(\zeta_t, X_t)\exp\{K_t + \frac{1}{2} < K >_t\} \,|\, \mathcal{Y}_t]. \tag{11}$$

But from (4)-(5), (6)-(7) and the definition of $\widetilde{K}_t(z)$ in (9), we can write

$$K_t = \widetilde{K}_t - < K >_t,$$

where

$$\widetilde{K}_t \triangleq \widetilde{K}_t(\zeta). \tag{12}$$

Therefore equation (11) may be rewritten as

$$\widetilde{\mathbb{E}}[\psi(\zeta_t, X_t)L_t^{-1} \,|\, \mathcal{Y}_t] = \widetilde{\mathbb{E}}[\psi(\zeta_t, X_t)\exp\{\widetilde{K}_t - \frac{1}{2} < K >_t\} \,|\, \mathcal{Y}_t].$$

Now since we assume that ζ and X_0 are independent, ζ is also independent of (X_0, V, W) under $\widetilde{\mathbb{P}}$. Hence, due to (9), ζ and $(\widetilde{V}, \widetilde{W}, X, Y)$ are also independent under $\widetilde{\mathbb{P}}$. Since moreover the distribution of ζ under $\widetilde{\mathbb{P}}$ is the same as under $\mathbb{P}$, taking into account the definition of $< K >_t$ in (6) and the definition (12) of $\widetilde{K}_t$, the representation (10) is readily obtained from the last equation. $\square$

Remark 2.1 Let us mention that if V is a non standard Brownian motion ($\mathbb{E}[V_1 V_1'] = \Sigma_V$) then, provided that the condition

$$(I_d - \Sigma_V \Sigma_V^+) b_\zeta \equiv 0,$$

where Σ_V^+ denotes a pseudo-inverse of Σ_V, is fulfilled, all which has just been discussed remains valid by replacing the definitions (4)-(5) by

$$K_t(z) \triangleq \int_0^t b_\zeta(z_s)' \Sigma_V^+ dV_s + \int_0^t h_\zeta(z_s)' dW_s, \; t \geq 0, \tag{13}$$

and

$$< K >_t (z) \triangleq \int_0^t b_\zeta(z_s)' \Sigma_V^+ b_\zeta(z_s) ds + \int_0^t |h_\zeta(z_s)|^2 ds, \; t \geq 0. \tag{14}$$

In particular the statement in Lemma 2.1 holds without any change, but of course in that case the process $\widetilde{V}$ defined in (7) and appearing in (9) is a Brownian motion such that $\widetilde{\mathbb{E}}[\widetilde{V}_1 \widetilde{V}_1'] = \Sigma_V$.

2.2 Linear filtering and Kalman filters.

Here we discuss a particular case where all functions b, b_ζ, h and h_ζ in model (1)-(2) are linear. We assume that for some matrices B, B_ζ, H and H_ζ of appropriate dimensions and all $x \in \mathbb{R}^d,\ z \in \mathbb{R}^k$ we have

- $b(x) = Bx;\ b_\zeta(z) = B_\zeta z;\ h(x) = Hx;\ h_\zeta(z) = H_\zeta z.$

Moreover we assume that the process ζ is governed by the ordinary linear differential equation

$$\dot{\zeta}_t = A\zeta_t,\ t \geq 0,\ \zeta_0,$$

where ζ_0 stands for some random $\mathcal{F}_0$-measurable initial condition such that $\mathbb{E}[|\zeta_0|^2]$
$< +\infty$. Then (ζ, X_0) is independent of (V, W). It is clear that assumptions $(A_0) - (A_1) - (A_2)$ are fulfilled and actually Ω_0 may be reduced to $\mathcal{C}(\mathbb{R}^+; \mathbb{R}^k)$. The process (ζ, X, Y) taken under $\mathbb{P}$ is the solution to the following linear filtering model:

$$\left\{ \begin{array}{rcl} d\zeta_t &=& A\zeta_t dt,\ t \geq 0,\ \zeta_0, \\ dX_t &=& BX_t\, dt + B_\zeta \zeta_t dt + dV_t,\ t \geq 0,\ X_0, \\ dY_t &=& HX_t\, dt + H_\zeta \zeta_t dt + dW_t,\ t \geq 0,\ Y_0 = 0. \end{array} \right. \tag{15}$$

Obviously under the probability $\mathbb{P}^0$, which here corresponds to the initial condition $\zeta_0 = 0$, the system may be reduced to

$$\left\{ \begin{array}{rcl} dX_t &=& BX_t\, dt + dV_t,\ t \geq 0,\ X_0, \\ dY_t &=& HX_t\, dt + dW_t,\ t \geq 0,\ Y_0 = 0. \end{array} \right. \tag{16}$$

It is well known that if the pair (ζ_0, X_0) is Gaussian then both the exact filter Π_t and the basic approximate filter Π_t^0 are Kalman filters. Of course the Kalman filtering equations governing the mean and covariance for Π_t (resp. Π_t^0) are in dimensions $k + d$ and $(k + d) \times (k + d)$ (resp. d and $d \times d$). If the initial condition is non-Gaussian, Π_t and Π_t^0 are not Kalman filters but they can be expressed in terms of appropriate modified Kalman filters (see Makowski [7] and Beněs and Karatzas [1]).

In the sequel the notation $\mathcal{N}(\mu, \Lambda)$ will be used for the Gaussian law on some space $\mathbb{R}^l$ with mean $\mu \in \mathbb{R}^l$ and a symmetric nonnegative definite $l \times l$-matrix matrix Λ $(\Lambda \geq 0)$ as the covariance. Moreover for every $\varphi \in \mathcal{M}_b(\mathbb{R}^l)$ we shall write:

$$\mathcal{N}(\mu, \Lambda)(\varphi) = \int_{\mathbb{R}^l} \varphi(u)\, \mathcal{N}(\mu, \Lambda)(du).$$

For any vector $x \in \mathbb{R}^d$ and any $d \times d$ matrix $R \geq 0$, let $\hat{X}_t^{x,R}$ and Q_t^R be the solutions to the Kalman filtering equations corresponding to model (16) initialized with (x, R) i.e.

$$d\hat{X}_t^{x,R} = B\hat{X}_t^{x,R}\, dt + Q_t^R H' [dY_t - H\hat{X}_t^{x,R} dt]\,,\ t \geq 0,\ x, \tag{17}$$

$$\dot{Q}_t^R = BQ_t^R + Q_t^R B' + \Sigma_V - Q_t^R H' H Q_t^R\,,\ t \geq 0,\ R. \tag{18}$$

Then the Gaussian distribution

$$\Pi_t^{x,R} = \mathcal{N}(\hat{X}_t^{x,R}, Q_t^R),$$

will be referred to as an approximate Kalman filter initialized with (x, R). Of course if $\nu_0 = \mathcal{N}(m_0, Q_0)$, for the basic approximate filter Π_t^0, we have

$$\Pi_t^0 = \mathcal{N}(\hat{X}_t^{m_0, Q_0}, Q_t^{Q_0}).$$

By using the results of Ocone and Pardoux [8] concerning the asymptotic stability of optimal filters with respect to perturbations to their initial condition in linear systems, we can prove under some additional conditions that, whatever is the distribution of (ζ_0, X_0) provided that ζ_0 and X_0 have finite second order moments, any approximate Kalman filter is asymptotically optimal in the following sense:

Proposition 2.1 *Assume that in model (15), the matrix A is stable, the pair $(B, \Sigma_V^{\frac{1}{2}})$ is stabilizable and the pair (B, H) is detectable. Suppose also that $\mathbb{E}[|\zeta_0|^2]$ $< +\infty$ and $\mathbb{E}[|X_0|^2] < +\infty$. Then for any $x \in \mathbb{R}^d$ and any $d \times d$ matrix $R \geq 0$*

$$\lim_{t \to \infty} \{\hat{X}_t - \hat{X}_t^{x,R}\} = 0 \ \ \mathbb{P}\ a.s., \tag{19}$$

and in the $L^2(\mathbb{P})$ sense. Moreover

$$\lim_{t \to \infty} \{\Pi_t(\varphi) - \Pi_t^{x,R}(\varphi)\} = 0 \ \ \mathbb{P}\ a.s., \tag{20}$$

for every uniformly continuous function $\varphi \in \mathcal{M}_b(\mathbb{R}^d)$.

Proof Obviously, introducing $Z_t = (\zeta_t', X_t')'$, the system (15) can be rewritten as

$$\left\{ \begin{array}{rcl} dZ_t & = & \check{B} Z_t\, dt + d\check{V}_t, \ t \geq 0, \ Z_0, \\ dY_t & = & \check{H} Z_t\, dt + dW_t, \ t \geq 0, \ Y_0 = 0, \end{array} \right. \tag{21}$$

where

$$\check{B} = \left(\begin{array}{cc} A & O \\ B & B_\zeta \end{array} \right) ; \check{H} = (H | H_\zeta) \ ,$$

and $\check{V}_t = (0, V_t')'$, $t \geq 0$. Clearly $\check{V} = (\check{V}_t, t \geq 0)$ is a Brownian motion independent of W and

$$\Sigma_{\check{V}} = \mathbb{E}[\check{V}_1 \check{V}_1'] = \left(\begin{array}{cc} O & O \\ O & \Sigma_V \end{array} \right).$$

Moreover, for initial conditions z and $\check{R}$ of the form

$$z = (0, x')' ; \quad \check{R} = \left(\begin{array}{cc} O & O \\ O & R \end{array} \right),$$

the Kalman filtering equations for Z_t may be reduced to equations (17) and (18) initialized with (x, R). Finally it is easy to check that assumptions concerning matrices A, B, H and $\Sigma_V^{\frac{1}{2}}$ imply that the pair $(\check{B}, \Sigma_{\check{V}}^{\frac{1}{2}})$ is stabilizable and the pair $(\check{B}, \check{H})$ is detectable. Hence the assertions in the statement are immediate consequences of [8, Theorem 2.6]. $\square$

Remark 2.2 (a) Since the pairs $(B, \Sigma_V^{\frac{1}{2}})$ and (B, H) are stabilizable and detectable respectively, it is known (cf. e.g. [5]) that for any matrix $R \geq 0$ the solution Q_t^R of the differential Riccati equation (18) initialized with R converges as t tends to infinity to the unique symmetric nonnegative definite solution Q_∞ of the algebraic Riccati equation

$$BQ + QB' + \Sigma_V - QH'HQ = 0. \tag{22}$$

Then Proposition 2.1 says in particular that for any $x \in \mathbb{R}^d$ the approximate Kalman filter $\Pi_t^{x,Q_\infty} = \mathcal{N}(\hat{X}_t^{x,Q_\infty}, Q_\infty)$ is asymptotically optimal. (b) Let us mention that Proposition 2.1 can be interpreted in terms of filters of reduced dimensions concerning general linear systems (cf. [6]).

3 The semi-linear case.

In this section we extend the results of Section 2.2 to the case of what we call a semi-linear system. Here we assume that in model (1)-(2) the condition (A_0) is fulfilled and conditions $(A_1) - (A_2)$ are strengthened respectively into

- (A_1^*) $\quad b(x) = Bx,\ h(x) = Hx,\ x \in \mathbb{R}^d$; b_ζ and h_ζ are Lipschitzian functions,
- (A_2^*) $\quad$ for all $t \geq 0$ $\ \mathbb{E}[\int_0^t |\zeta_s|^2 ds] < +\infty;\ \lim_{t \to +\infty} \mathbb{E}[|\zeta_t|^2] = 0.$

Moreover we introduce the following assumption:

- (A_3) $\quad (B, \Sigma_V^{\frac{1}{2}})$ is a stabilizable pair; (B, H) is a detectable pair.

Then the process (X, Y) taken under $\mathbb{P}$ is the solution to the filtering model

$$\left\{ \begin{array}{rcl} dX_t & = & BX_t dt + b_\zeta(\zeta_t)dt + dV_t,\ t \geq 0,\ X_0, \\ dY_t & = & HX_t dt + h_\zeta(\zeta_t)dt + dW_t,\ t \geq 0,\ Y_0 = 0. \end{array} \right. \tag{23}$$

Obviously under the probability $\mathbb{P}^0$, which corresponds to $\zeta \equiv 0$, the system may be reduced to the linear system (16). Therefore again if $\nu_0 = \mathcal{N}(m_0, Q_0)$, the basic approximate filter Π_t^0 in model (23) is the Kalman filter defined as above by (17)-(18) initialized with (m_0, Q_0). Actually whatever is the prior distribution ν_0, all Kalman filters $\Pi_t^{x,R} = \mathcal{N}(\hat{X}_t^{x,R}, Q_t^R)$, $x \in \mathbb{R}^d,\ R \geq 0$ are candidates as approximate filters. The asymptotic study of these filters will rely on a representation of the exact filter Π_t which we derive now.

3.1 Representation of the exact filter.

The approach is similar to that used by Ocone and Pardoux [8] to study the case of a linear system with a non-Gaussian initial condition. >From (23) we may decompose the signal X as

$$X_t = \zeta_t^* + X_t^*,\ t \geq 0, \tag{24}$$

where

$$\zeta_t^* = \zeta_t^*(\zeta, X_0); \quad X_t^* = \int_0^t e^{B(t-s)} dV_s,\ t \geq 0,$$

with, for any $z = (z_t, t \geq 0) \in \Omega_0$ and $x_0 \in \mathbb{R}^d$,

$$\zeta_t^*(z, x_0) = e^{Bt}x_0 + \int_0^t e^{B(t-s)} b_\zeta(z_s) ds, \ t \geq 0. \tag{25}$$

Then the process (X^*, Y) appears as the solution of the system

$$\begin{cases} dX_t^* &= BX_t^* dt + dV_t, \ t \geq 0, \ X_0^* = 0, \\ dY_t &= HX_t^* dt + [H_\zeta \zeta_t^* + h_\zeta(\zeta_t)]dt + dW_t, \ t \geq 0, \ Y_0 = 0. \end{cases} \tag{26}$$

It can be seen as a version of model (1)-(2) with X^* in place of X and (ζ, ζ^*) in place of ζ. Notice that since (ζ, X_0) is independent of (V, W) the pair (ζ, ζ^*) is also independent of (V, W). Hence we may apply the considerations in Section 2.1. So let us define for $x_0 \in \mathbb{R}^d$, $z = (z_t, t \geq 0) \in \Omega_0$ and $t \geq 0$

$$K_t^*(z, x_0) \triangleq \int_0^t [H\zeta_s^*(z, x_0) + h_\zeta(z_s)]' dW_s, \tag{27}$$

$$< K >_t^* (z, x_0) \triangleq \int_0^t |H\zeta_s^*(z, x_0) + h_\zeta(z_s)|^2 ds, \tag{28}$$

and

$$K_t^* \triangleq K_t^*(\zeta, X_0); \ < K >_t^* \triangleq < K >_t^* (\zeta, X_0); \ L_t^* \triangleq \exp\{-K_t^* - \frac{1}{2} < K >_t^*\}.$$

Let $\mathbb{P}^*$ be the probability on Ω which is locally absolutely continuous with respect to $\mathbb{P}$ with the local Radon-Nikodym derivative L_t^*. Let also $\mathbb{E}^*[\,.\,]$ denote an expectation with respect to $\mathbb{P}^*$. Then by (24) and the Bayes formula we get the following representation of the exact filter:

$$\Pi_t(\varphi) = \frac{\mathbb{E}^*[\varphi(\zeta_t^* + X_t^*)(L_t^*)^{-1} \mid \mathcal{Y}_t]}{\mathbb{E}^*[(L_t^*)^{-1} \mid \mathcal{Y}_t]}, \ \ \varphi \in \mathcal{M}^+(\mathbb{R}^d). \tag{29}$$

Moreover, since here $X_0^* = 0$ and (ζ, ζ^*) is independent of (V, W) and consequently (X_0^*, ζ, ζ^*) is independent of (V, W), we may apply Lemma 2.1 (see also Remark 2.1). Therefore we obtain that for any function $\psi \in \mathcal{M}^+(\mathbb{R}^{k+d})$ we have

$$\mathbb{E}^*[\psi(\zeta_t^*, X_t^*)(L_t^*)^{-1} \mid \mathcal{Y}_t] = \int_{\Omega_0 \times \mathbb{R}^d} e^{-\frac{1}{2}<K>_t^*(z,x_0)} I_t(z, x_0) d\mathbb{P}_{(\zeta, X_0)}(z, x_0), \tag{30}$$

with

$$I_t(z,x_0) = \int_{\mathbb{R}^{d+1}} \psi(\zeta_t^*(z,x_0),u)e^v n_t^{z,x_0}(du,dv), \tag{31}$$

where for all $z=(z_t, t\geq 0)\in\Omega_0$, $x_0\in\mathbb{R}^d$ and $t\geq 0$ the symbol n_t^{z,x_0} stands for the conditional distribution of $(X_t^*, K_t^*(z,x_0))$ given $\mathcal{Y}_t$ in the model

$$\left\{\begin{array}{lcl} dX_t^* & = & BX_t^*dt + dV_t;\ t\geq 0;\ X_0^*=0, \\ dK_t^*(z,x_0) & = & [H\zeta_t^*(z,x_0)+h_\zeta(z_t)]'dW_t^*;\ t\geq 0;\ K_0^*(z,x_0)=0, \\ dY_t & = & HX_t^*dt + dW_t^*;\ t\geq 0;\ Y=0. \end{array}\right. \tag{32}$$

Here of course (32) is taken under the probability $\mathbb{P}^*$ with respect to which the process $W^*=(W_t^*, t\geq 0)$ defined by

$$W_t^* \triangleq \int_0^t [H\zeta_s^* + h_\zeta(\zeta_s)]ds + W_t,\ t\geq 0,$$

is a standard Brownian motion independent of V.
Let us introduce $Z_t(z,x_0)=(X_t^{*\prime}, K_t^*(z,x_0))'$,

$$B_t(z,x_0) = \left(\begin{array}{cc} B & O \\ -[H\zeta_t^*(z,x_0)+h_\zeta(z_t)]'H & O \end{array}\right); G_t(z,x_0) = \left(\begin{array}{c} O \\ [H\zeta_t^*(z,x_0)+h_\zeta(z_t)]' \end{array}\right),$$

and $H^*=(H|O)$. Define also $V_t^*=(V_t',0)'$, $t\geq 0$ such that, under $\mathbb{P}^*$, the process $V^*=(V_t^*,\ t\geq 0)$ is a Brownian motion independent of W^* and

$$\Sigma_{V^*} = \mathbb{E}^*[V_1^*V_1^{*\prime}] = \left(\begin{array}{cc} \Sigma_V & O \\ O & O \end{array}\right).$$

Then the system (32) appears as the following conditionally linear Gaussian system for $(Z_t(z,x_0),Y)$

$$\left\{\begin{array}{lcl} dZ_t & = & B_tZ_tdt + G_tdY_t + dV_t^*,\ t\geq 0,\ Z_0=0, \\ dY_t & = & H^*Z_tdt + dW_t^*,\ t\geq 0,\ Y_0=0, \end{array}\right.$$

where the dependences on (z,x_0) are omitted as it will be done often from now on. Hence the conditional distribution n_t^{z,x_0} is nothing but

$$n_t^{z,x_0} = \mathcal{N}(\hat{Z}_t(z,x_0), C_t(z,x_0)),$$

where the mean $\hat{Z}_t(z,x_0)$ and the covariance $C_t(z,x_0)$ are governed by the Kalman filtering equations

$$\left\{\begin{array}{lcl} d\hat{Z}_t & = & B_t\hat{Z}_t dt + G_t dY_t + C_t {H^*}^{'}[dY_t - H^*\hat{Z}_t dt],\ t \geq 0,\ \hat{Z}_0 = 0, \\ \dot{C}_t & = & B_t C_t + C_t B_t^{'} + \Sigma_{V^*} - C_t {H^*}^{'} H^* C_t,\ t \geq 0,\ C_t = 0. \end{array}\right. \tag{33}$$

Let

$$\hat{Z}_t \triangleq \left(\begin{array}{c} \hat{X}_t^* \\ \hat{K}_t^* \end{array}\right); \quad C_t \triangleq \left(\begin{array}{cc} Q_t^* & S_t \\ S_t^{'} & T_t \end{array}\right), \tag{34}$$

where $\hat{X}_t^* \in \mathbb{R}^d$, $\hat{K}_t^* \in \mathbb{R}$, $Q_t^* \in \mathbb{R}^{d\times d}$, $S_t \in \mathbb{R}^d$ and $T_t \in \mathbb{R}$. Then from (33) we have

$$d\hat{X}_t^* = B\hat{X}_t^* dt + Q_t^* H^{'}[dY_t - H\hat{X}_t^* dt],\ t \geq 0,\ \hat{X}_0^* = 0, \tag{35}$$

$$\dot{Q}_t^* = BQ_t^* + Q_t^* B^{'} + \Sigma_V - Q_t^* H^{'} H Q_t^*,\ t \geq 0,\ Q_0^* = O\,, \tag{36}$$

$$\dot{S}_t = (B - Q_t^* H^{'} H) S_t - Q_t^* H^{'}[H\zeta_t^* + h_\zeta(\zeta_t)],\ t \geq 0,\ S_0 = 0\,, \tag{37}$$

$$\dot{T}_t = -2[H\zeta_t^* + h_\zeta(z_t)]^{'} H S_t - S_t^{'} H^{'} H S_t,\ t \geq 0,\ T_0 = 0. \tag{38}$$

Remark 3.1 Comparing equations (35)-(36) to (17)-(18), it is clear that $\hat{X}_t^* = \hat{X}_t^{0,O}$ and $Q_t^* = \hat{Q}_t^O$, where $\hat{X}_t^{0,O}$ and Q_t^O are the solutions of (17) and (18) corresponding to initial conditions 0 and O respectively. In other words the marginal distribution of n_t^{z,x_0} on $\mathbb{R}^d$ is nothing but the approximate Kalman filter $\Pi_t^{0,O}$ for X_t, defined by (17)-(18), initialized with $(0, O)$.

Now we are able to prove

Proposition 3.1 *Let Π_t be the exact filter in model (23) (taken under $\mathbb{P}$). Then for any $\varphi \in \mathcal{M}^+(\mathbb{R}^d)$ the following holds:*

$$\Pi_t(\varphi) = \frac{\int_{\Omega_0\times\mathbb{R}^d} e^{[-\frac{1}{2}<K>_t^* + \frac{1}{2}T_t + \hat{K}_t^*](z,x_0)} J_t(z,x_0)\, d\mathbb{P}_{(\zeta,X_0)}(z,x_0)}{\int_{\Omega_0\times\mathbb{R}^d} e^{[-\frac{1}{2}<K>_t^* + \frac{1}{2}T_t + \hat{K}_t^*](z,x_0)} d\mathbb{P}_{(\zeta,X_0)}(z,x_0)}, \tag{39}$$

with

$$J_t(z,x_0) = \int_{\mathbb{R}^{d+1}} \varphi(\zeta_t^*(z,x_0) + u)\, \bar{n}_t^{z,x_0}(du,dv), \tag{40}$$

where for all $z=(z_t, t\geq 0)\in\Omega_0$, $x_0\in I\!R^d$ and $t\geq 0$ the symbol $\bar{n}_t^{z,x_0}$ stands for the Gaussian distribution on $I\!R^{d+1}$ with mean $([\hat{X}_t^*+S_t(z,x_0)]', \hat{K}_t^*(z,x_0)+T_t(z,x_0))'$ and covariance $C_t(z,x_0)$. Here the quantities depending on (z,x_0) are defined by (25), (27) and (34)-(38).

Proof From (29)-(30)-(31) we get

$$\Pi_t(\varphi)=\frac{\int_{\Omega_0\times I\!R^d} e^{-\frac{1}{2}<K>_t^*(z,x_0)} I_t^{(1)}(z,x_0)d\mathbb{P}_{(\zeta,X_0)}(z,x_0)}{\int_{\Omega_0\times I\!R^d} e^{-\frac{1}{2}<K>_t^*(z,x_0)} I_t^{(2)}(z,x_0)d\mathbb{P}_{(\zeta,X_0)}(z,x_0)},$$

with

$$\begin{aligned} I_t^{(1)}(z,x_0) &= \int_{I\!R^{d+1}} \varphi(\zeta_t^*(z,x_0)+u)e^v n_t^{z,x_0}(du,dv),\\ I_t^{(2)}(z,x_0) &= \int_{I\!R^{d+1}} e^v n_t^{z,x_0}(du,dv), \end{aligned}$$

where $n_t^{z,x_0}=\mathcal{N}(\hat{Z}_t(z,x_0),C_t(z,x_0))$. In order to get (39)-(40) it remains to show that the probability distribution $\bar{n}_t^{z,x_0}$ defined by

$$\bar{n}_t^{z,x_0}(du,dv)\triangleq e^{[-\frac{1}{2}T_t-\hat{K}_t^*](z,x_0)}e^v n_t^{z,x_0}(du,dv),$$

is nothing but the Gaussian distribution on $\mathbb{R}^{d+1}$ with mean $([\hat{X}_t^*+S_t(z,x_0)]', \hat{K}_t^*(z,x_0)+T_t(z,x_0))'$ and covariance $C_t(z,x_0)$. To do that we compute the characteristic functional of $\bar{n}_t^{z,x_0}$. First we notice that for any $\mu\in\mathbb{R}^d$ and $\lambda\in\mathbb{R}$

$$\begin{aligned} &\int_{I\!R^{d+1}} e^{i<\lambda,u>+i\mu v+v} n_t^{z,x_0}(du,dv)\\ =\ & e^{i<\lambda,\hat{X}_t^*>+i(\mu-i)\hat{K}_t^*}\exp\left\{-\frac{1}{2}\left\langle C_t\begin{pmatrix}\lambda\\ \mu-i\end{pmatrix},\begin{pmatrix}\lambda\\ \mu-i\end{pmatrix}\right\rangle\right\}\\ =\ & e^{\frac{1}{2}T_t+\hat{K}_t^*}\exp\left\{i<\lambda,\hat{X}_t^*+S_t>+i\mu(\hat{K}_t^*+T_t)-\frac{1}{2}\left\langle C_t\begin{pmatrix}\lambda\\ \mu\end{pmatrix},\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\right\rangle\right\}. \end{aligned}$$

Then we obtain

$$\begin{aligned} &\int_{I\!R^{d+1}} e^{i<\lambda,u>+i\mu v}\,\bar{n}_t^{z,x_0}(du,dv)\\ =\ & \exp\left\{i<\lambda,\hat{X}_t^*+S_t>+i\,\mu(\hat{K}_t^*+T_t)-\frac{1}{2}\left\langle C_t\begin{pmatrix}\lambda\\ \mu\end{pmatrix},\begin{pmatrix}\lambda\\ \mu\end{pmatrix}\right\rangle\right\}, \end{aligned}$$

which gives the result. □

Remark 3.2 Notice that substituting $\mathbb{P}^0$ for $\mathbb{P}$ in Proposition 3.1, i.e. taking $\zeta \equiv 0$ and then also $\zeta_t^*(z, x_0) \equiv e^{Bt}x_0$, gives a representation of the approximate filter Π_t^0. Of course the corresponding formula reduces to that obtained by Makowski [7] and Bens and Karatzas [1] for the optimal filter in a linear system with non-Gaussian initial condition.

3.2 Asymptotic optimality of the approximate filter.

Now we are able to show that an approximate Kalman filter $\Pi_t^{x,R} = \mathcal{N}(\hat{X}_t^{x,R}, Q_t^R)$ defined by (17)-(18) and initialized with arbitrary $x \in \mathbb{R}^d$ and $R \geq 0$ is asymptotically optimal when $\mathbb{E}[|X_0|^2] < +\infty$. At first we state a result concerning the approximation of the conditional mean $\hat{X}_t$ by the solution $\hat{X}_t^{x,R}$ of (17)-(18).

Proposition 3.2 *Assume that in model (23), the conditions* $(A_0) - (A_1^*) - (A_2^*) - (A_3)$ *are fulfilled. Suppose also that* $\mathbb{E}[|X_0|^2] < +\infty$. *Then for any* $x \in \mathbb{R}^d$ *and any* $d \times d$ *matrix* $R \geq 0$

$$\lim_{t \to \infty} \mathbb{E}[|\hat{X}_t - \hat{X}_t^{x,R}|^2] = 0. \tag{41}$$

Proof Since $\mathbb{E}[|X_0|^2] < +\infty$, from assumptions $(A_1^*) - (A_2^*)$ it is easy to check that ζ_t^* and X_t^* are integrable. Then the equalities (39)-(40) are satisfied in the case $\varphi(x) = x, x \in \mathbb{R}^d$. In this case we have

$$\int_{\mathbb{R}^{d+1}} \varphi\left(\zeta_t^*(z, x_0) + u\right) \bar{n}_t^{z,x_0}(du, dv) = \zeta_t^*(z, x_0) + \int_{\mathbb{R}^d} u\bar{m}_t^{z,x_0}(du),$$

where $\bar{m}_t^{z,x_0}$ is the marginal of $\bar{n}_t^{z,x_0}$ on $\mathbb{R}^d$ i.e. $\bar{m}_t^{z,x_0}$ is the Gaussian law with mean $\hat{X}_t^* + S_t(z, x_0)$ and covariance Q_t^*. Therefore

$$\int_{\mathbb{R}^{d+1}} \varphi\left(\zeta_t^*(z, x_0) + u\right) \bar{n}_t^{z,x_0}(du, dv) = \zeta_t^*(z, x_0) + \hat{X}_t^* + S_t(z, x_0). \tag{42}$$

>From (24), (39)-(40) and (42) we get

$$\hat{X}_t = \hat{X}_t^* + \mathbb{E}[\zeta_t^* + S_t \,|\, \mathcal{Y}_t]\,. \tag{43}$$

Let us consider $\mathbb{E}[\zeta_t^* + S_t \,|\, \mathcal{Y}_t]$. From equations (25) and (37) we obtain

$$d(\zeta_t^* + S_t) = (B - Q_t^* H' H)(\zeta_t^* + S_t)\, dt + [b_\zeta(\zeta_t) - Q_t^* H' h_\zeta(\zeta_t)]\, dt\,, \qquad (44)$$

with $\zeta_0^* + S_0 = X_0$. Thus

$$\zeta_t^* + S_t = \Phi_t X_0 + \Phi_t \int_0^t \Phi_s^{-1}[b_\zeta(\zeta_s) - Q_s^* H' h_\zeta(\zeta_s)]\, ds\,, \qquad (45)$$

where $\Phi_t \in \mathbb{R}^{d\times d}$ is the transition matrix associated with the matrix $B - Q_t^* H' H$. Under the assumption (A_3) it is known (cf. e.g. [5] and [8]) that the solution Q_t^* of equation (36) converges as t goes to $+\infty$ to the unique solution Q_∞ of the Riccati algebraic equation (22) which is such that the matrix $B - Q_\infty H' H$ is asymptotically stable. Moreover, if λ_1 denotes the maximum of the real parts of eigenvalues of $B - Q_\infty H' H$, for every $0 < \sigma < -\lambda_1$, there exist a constant $C_\sigma > 0$ and a $t_\sigma < +\infty$ such that

$$||\Phi_t \Phi_s^{-1}|| \le C_\sigma e^{-\sigma(t-s)}\,, \text{ for } t_\sigma \le s < t\,, \qquad (46)$$

and

$$||Q_t^* - Q_\infty|| \le C_\sigma e^{-\sigma t}\,, \text{ for all } t \ge 0\,. \qquad (47)$$

Then we rewrite the equation (45) for $t \ge t_\sigma$ as

$$\begin{aligned} \zeta_t^* + S_t \;=\; & \Phi_t \Phi_{t_\sigma}^{-1}(\zeta_{t_\sigma}^* + S_{t_\sigma}) \\ & + \Phi_t \textstyle\int_{t_\sigma}^t \Phi_s^{-1}[b_\zeta(\zeta_s) - Q_\infty H' h_\zeta(\zeta_s)]\, ds \qquad (48) \\ & + \Phi_t \textstyle\int_{t_\sigma}^t \Phi_s^{-1}(Q_\infty - Q_s^*) H' h_\zeta(\zeta_s)\, ds\,. \end{aligned}$$

Hence, making use of (46)-(47), by assumptions (A_1^*) we have

$$\begin{aligned} &\mathbb{E}[|\zeta_t^* + S_t|^2] \le 4\mathbb{E}[|\Phi_t \Phi_{t_\sigma}^{-1}(\zeta_{t_\sigma}^* + S_{t_\sigma})|^2] \\ &+ \; 4C_\sigma \int_{t_\sigma}^t e^{-2\sigma(t-s)} \mathbb{E}[|\zeta_s|^2]\, ds + 4C_\sigma \int_{t_\sigma}^t e^{-2\sigma(t-s)} e^{-2\sigma s} \mathbb{E}[|\zeta_s|^2]\, ds\,. \end{aligned}$$

Then there exist a constant $C_\sigma > 0$ and a $t_\sigma < \infty$ such that for $t \ge t_\sigma$

$$\mathbb{E}[|\zeta_t^* + S_t|^2] \le C_\sigma \left[e^{-2\sigma(t-t_\sigma)} + \int_{t_\sigma}^t e^{-2\sigma(t-s)} \mathbb{E}[|\zeta_s|]^2\, ds \right]\,. \qquad (49)$$

Taking assumption (A_2^*) into account and making use of the Tœplitz Lemma we obtain

$$\lim_{t \to \infty} \mathbb{E}[|\zeta_t^* + S_t|^2] = 0\,. \tag{50}$$

Since $\hat{X}_t - \hat{X}_t^* = \mathbb{E}(\zeta_t^* + S_t \,|\, \mathcal{Y}_t)$, by the Jensen inequality, it follows that

$$\lim_{t \to \infty} \mathbb{E}[|\hat{X}_t - \hat{X}_t^*|^2] = 0\,. \tag{51}$$

To complete the proof of (41), we have to show that $\hat{X}_t^* - \hat{X}_t^{x,R}$ converges to 0 in $L^2(\mathbb{P})$ for any $x \in \mathbb{R}^d$ and $R \geq 0$. More precisely, we are going to show that for any $0 < \sigma < -\lambda_1$ there exist a constant C_σ and a $t_\sigma < \infty$ such that

$$\mathbb{E}[|\hat{X}_t^* - \hat{X}_t^{x,R}|^2] \leq C_\sigma e^{-2\sigma t}\,, \quad \text{for all } t \geq 0\,. \tag{52}$$

Recall (cf. Remark 3.1) that $\hat{X}_t^* = \hat{X}_t^{0,O}$ and $Q_t^* = Q^O$. Then from equations (17) and (35) we obtain

$$\begin{aligned} d(\hat{X}_t^* - \hat{X}_t^{x,R}) &= (B - Q_t^R H'H)(\hat{X}_t^* - \hat{X}_t^{x,R})\,dt + (Q_t^* - Q_t^R)H'\,d\nu_t \\ &\quad + (Q_t^* - Q_t^R)H'H(\hat{X}_t - \hat{X}_t^*)\,dt \\ &\quad + (Q_t^* - Q_t^R)H'\widehat{h_\zeta(\zeta_t)}\,dt, \end{aligned}$$

where $\nu_t = Y_t - \int_0^t (H\hat{X}_s + \widehat{h_\zeta(\zeta_s)})\,ds$ defines the innovation process and the initial condition is $\hat{X}_0^* - \hat{X}_0^{x,R} = -x$. Then for a given $t_\sigma > 0$, we have

$$\begin{aligned} \hat{X}_t^* - \hat{X}_t^{x,R} &= \Phi_t^R(\Phi_{t_\sigma}^R)^{-1}(\hat{X}_{t_\sigma} - \hat{X}_{t_\sigma}^{x,R}) + \Phi_t^R \int_{t_\sigma}^t (\Phi_s^R)^{-1}(Q_s^* - Q_s^R)H'\,d\nu_s \\ &\quad + \Phi_t^R \int_{t_\sigma}^t (\Phi_s^R)^{-1}(Q_s^* - Q_s^R)H'H(\hat{X}_s - \hat{X}_s^*)ds \\ &\quad + \Phi_t^R \int_{t_\sigma}^t (\Phi_s^R)^{-1}(Q_s^* - Q_s^R)H'\widehat{h_\zeta(\zeta_s)}\,ds\,, \end{aligned}$$

where Φ_t^R is the transition matrix associated with $B - Q_t^R H'H$.
To prove (52) it suffices to establish the bound for each of the terms in the above right-hand side. The first two terms can be treated as in (48) and the bound for the last two terms is easily obtained. This completes the proof of (41). $\square$

Now we state a result concerning the approximation of the exact filter Π_t by a Kalman filter $\Pi_t^{x,R} = \mathcal{N}(\hat{X}_t^{x,R}, Q_t^R)$ defined by (17)-(18).

Proposition 3.3 *Assume that in model (23), the conditions* $(A_0) - (A_1^*) - (A_2^*) - (A_3)$ *are fulfilled. Suppose also that* $\mathbb{E}[|X_0|^2] < +\infty$. *Then for any* $x \in \mathbb{R}^d$ *and any* $d \times d$ *matrix* $R \geq 0$

$$\lim_{t \to \infty} \mathbb{E}[|\Pi_t(\varphi) - \Pi_t^{x,R}(\varphi)|^2] = 0, \tag{53}$$

for all uniformly continuous $\varphi \in \mathcal{M}_b(\mathbb{R}^d)$.

Preparing for the proof let us first show

Lemma 3.1 *Let* $t \to x_t$ *and* $t \to \bar{x}_t$ *(resp.* $t \to Q_t$ *and* $t \to \bar{Q}_t$*) be continuous functions from* $\mathbb{R}^+$ *into* $\mathbb{R}^d$ *(resp. the set of symmetric nonnegative definite* $d{\times}d$ *matrices). Assume that* $\lim_{t \to \infty}(x_t - \bar{x}_t) = 0$ *and* $\lim_{t \to \infty} Q_t = Q_\infty = \lim_{t \to \infty} \bar{Q}_t$. *Then*

$$\lim_{t \to \infty} \left[\mathcal{N}(x_t, Q_t)(\varphi) - \mathcal{N}(\bar{x}_t, \bar{Q}_t)(\varphi)\right] = 0\,,$$

for every uniformly continuous function $\varphi \in \mathcal{M}_b(\mathbb{R}^d)$.

Proof Clearly,

$$\lim_{t \to \infty} \left[\mathcal{N}(0, Q_t)(\varphi) - \mathcal{N}(0, \bar{Q}_t)(\varphi)\right] = \lim_{t \to \infty} [\mathcal{N}(0, Q_t)(\varphi) - \mathcal{N}(0, Q_\infty)(\varphi)]$$
$$+ \lim_{t \to \infty} \left[\mathcal{N}(0, \bar{Q}_t)(\varphi) - \mathcal{N}(0, Q_\infty)(\varphi)\right] = 0\,.$$

Then on some probability space there exist random processes X and $\bar{X}$ taking values in $\mathrm{I\!R}^d$, such that X_t has law $\mathcal{N}(0, Q_t)$ and $\bar{X}_t$ has law $\mathcal{N}(0, \bar{Q}_t)$ for all $t \geq 0$, and moreover $|X_t - \bar{X}_t| \to 0$ almost surely. Let $U_t = X_t + x_t$ and $\bar{U}_t = \bar{X}_t + \bar{x}_t$. Since $\lim_{t \to \infty} x_t - \bar{x}_t = 0$, we obtain $|U_t - \bar{U}_t| \to 0$ almost surely. Given $\eta > 0$,

$$|\mathcal{N}(x_t, Q_t)(\varphi) - \mathcal{N}(\bar{x}_t, \bar{Q}_t)(\varphi)| = |\mathbf{E}\left(\varphi(U_t) - \varphi(\bar{U}_t)\right)|$$
$$\leq \sup_{|u-u'| \leq \eta} |\varphi(u) - \varphi(u')| + 2||\varphi||_\infty P(\{|U_t - \bar{U}_t| \geq \eta\})\,.$$

By first letting $t \to \infty$ and then using uniform continuity, we complete the proof. □

Proof of Proposition 3.3 >From (52), by applying the Borel-Cantelli lemma to $\hat{X}_n^* - \hat{X}_n^{x,R}$ and to $\sup_{n\le t\le n+1}|\hat{X}_t^* - \hat{X}_t^{x,R} - (\hat{X}_n^* - \hat{X}_n^{x,R})|$, we can easily obtain that $\hat{X}_t^* - \hat{X}_t^{x,R}$ converges to 0 almost surely. Since

$$\lim_{t\to\infty} Q_t^* = Q_\infty = \lim_{t\to\infty} Q_t^R\,,$$

according to Lemma 3.1 and the dominated convergence Theorem, we obtain

$$\lim_{t\to\infty}\mathbb{E}[|\mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi) - \mathcal{N}(\hat{X}_t^{x,R}, Q_t^R)(\varphi)|^2] = 0\,.$$

Therefore it remains to show that

$$\lim_{t\to\infty}\mathbb{E}[|\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi)|^2] = 0\,, \tag{54}$$

for any uniformly continuous function $\varphi\in\mathcal{M}_b(\mathbb{R}^d)$. Observe that

$$\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi) = \mathbb{E}[\varphi(X_t)\,|\,\mathcal{Y}_t] - \mathbb{E}[\varphi(\hat{X}_t^* + U_t)]\,,$$

where for each $t\ge 0$, U_t is a Gaussian vector with mean zero and covariance matrix Q_t^*.
From (39)-(40) we have

$$\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi) = \frac{\int_{\Omega_0\times\mathbb{R}^d} e^{\left(-\frac{1}{2}<K>_t+\frac{1}{2}T_t+\hat{K}_t^*\right)(z,x_0)}\Delta_t(z,x_0)\,d\mathbb{P}_{(\zeta,X_0)}(z,x_0)}{\int_{\Omega_0\times\mathbb{R}^d} e^{\left(-\frac{1}{2}<K>_t+\frac{1}{2}T_t+\hat{K}_t^*\right)(z,x_0)}d\mathbb{P}_{(\zeta,X_0)}(z,x_0)}\,, \tag{55}$$

where

$$\Delta_t(z,x_0) = \mathbb{E}[\varphi(\hat{X}_t^* + (S_t+\zeta_t^*)(z,x_0) + U_t)] - \mathbb{E}[\varphi(\hat{X}_t^* + U_t)]\,.$$

Let $\varepsilon>0$. Choose $\eta>0$ such that for all y and $y'\in\mathbb{R}^d$, if $|y-y'|\le\eta$ then

$$|\varphi(y)-\varphi(y')|^2 < \frac{\varepsilon}{4}\,.$$

Decompose the integral in the numerator of (55) into the sum of the integral over the region $\{|(\zeta_t^*+S_t)(z,x_0)|<\eta\}$ and the integral over the region $\{|(\zeta_t^*+S_t)(z,x_0)|\ge\eta\}$. We obtain

$$\begin{aligned}|\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi)| \;\le\; & \sup_{|y-y'|<\eta}|\varphi(y)-\varphi(y')| \\ & +2||\varphi||_\infty\mathbb{E}[\mathbb{1}_{\{|(\zeta_t^*+S_t)|\ge\eta\}}\,|\,\mathcal{Y}_t]\,,\end{aligned}$$

and therefore

$$\mathbb{E}[|\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi)|^2] \leq \frac{\varepsilon}{2} + \frac{8||\varphi||_\infty^2}{\eta^2}\mathbb{E}[|\zeta_t^* + S_t|^2].$$

Since by (50) $\lim_{t \to \infty} \mathbb{E}[|\zeta_t^* + S_t|^2] = 0$, there is a $T > 0$ such that for all $t > T$

$$\mathbb{E}[|\zeta_t^* + S_t|^2] \leq \frac{\eta^2}{16||\varphi||_\infty^2},$$

and hence also

$$\mathbb{E}[|\Pi_t(\varphi) - \mathcal{N}(\hat{X}_t^*, Q_t^*)(\varphi)|^2] \leq \varepsilon.$$

Thus (54) is obtained which achieves the proof of (53). □

Remark 3.3 One consequence of Propositions 3.2 and 3.3 is that the approximate filter is "asymptotically insensitive to perturbations of its initial condition". It is interesting to note that this result does not require asymptotic ergodic behavior of the signal. The statements extend to the case of colored noises in the signal and in the observation the one obtained by Ocone and Pardoux [8, Theorem 2.6] for white noises. Moreover these statements can be seen also as some extensions of Proposition 2.1 above to the case of a general nonlinear action of the perturbation process ζ on the basically linear dynamics of the signal-observation process.

4 The nonlinear case.

In this section we consider the case where in model (1)-(2) all functions b, b_ζ, h and h_ζ may be nonlinear but the system has a limiting ergodic behavior. Here we assume that the conditions (A_0) and (A_1') are fulfilled. Assumption (A_2) is strengthened into $(A_{2,1}^*) - (A_{2,2}^*)$ as follows. The process ζ is supposed to be generated by the ordinary differential equation

$$\dot{\zeta}_t = a(\zeta_t), \; t \geq 0, \; \zeta_0,$$

where a is a Lipschitzian function from $\mathbb{R}^k$ into $\mathbb{R}^k$. Concerning the deterministic flow $\phi = (\phi_t, t \geq 0)$ associated with that equation we suppose that

- $(A^*_{2,1})$ $\phi(0) \equiv 0$ and ϕ is contracting with some exponential rate $\lambda > 0$ i.e.
$$|\phi_t(z_1) - \phi_t(z_2)| \leq e^{-\lambda t}|z_1 - z_2|,$$
for all $t \geq 0$ and $z_1,\ z_2 \in \mathbb{R}^k$.

Concerning the initial condition ζ_0 we introduce the assumption:

- $(A^*_{2,2})$ ζ_0 is a $\mathcal{F}_0$-measurable initial condition such that
$$\mathbb{E}[\exp c_0|\zeta_0|^2] < +\infty,$$
for $c_0 = K_\zeta^2/\lambda$, where K_ζ is the maximum of the Lipschitz constants for b_ζ and h_ζ.

Moreover we make the following assumption on the drift coefficient b:

- (A^*_3) there exist two constants $r > 0$ and $\alpha > 0$ such that for $x \in \mathbb{R}^d$:
$$x'b(x) \leq -\alpha|x| \quad \text{if} \quad |x| > r.$$

Finally here we assume that the Brownian motion V is non-degenerate i.e. the matrix Σ_V is positive.
The concerned process (ζ, X, Y) taken under $\mathbb{P}$ on Ω (where actually Ω_0 may be reduced to $\mathcal{C}(\mathbb{R}^+;\mathbb{R}^k)$) is the $(\mathcal{F}_t)$-Markov diffusion process solution of the system
$$\left\{\begin{array}{rcl} d\zeta_t &=& a(\zeta_t)dt,\ t \geq 0,\ \zeta_0, \\ dX_t &=& b(X_t)dt + b_\zeta(\zeta_t)dt + dV_t,\ t \geq 0,\ X_0, \\ dY_t &=& h(X_t)dt + h_\zeta(\zeta_t)dt + dW_t,\ t \geq 0,\ Y_0 = 0. \end{array}\right. \tag{56}$$
The determination of the exact filter Π_t is, in general, an infinite dimensional problem. Actually the evolution in time of Π_t can be described by the so-called Zakai equation (cf. [10] or e.g. [9], [2]) where of course the spatial dimension is $k + d$ and the initial condition is $\mathbb{P}_{(\zeta_0, X_0)}$ which will be denoted by π_0.
Under the probability $\mathbb{P}^0$, which here corresponds to $\zeta_0 = 0$, the system may be reduced to
$$\left\{\begin{array}{rcl} dX_t &=& b(X_t)dt + dV_t,\ t \geq 0,\ X_0, \\ dY_t &=& h(X_t)dt + dW_t,\ t \geq 0,\ Y_0 = 0. \end{array}\right. \tag{57}$$

Then the evolution in time of the approximate filter Π_t^0 is described by a Zakai equation for which numerical computations are easier since the spatial dimension is d only and the initial condition is the distribution ν_0 of X_0.
Here the asymptotic study of the approximate filter will use results of Ocone and Pardoux [8] on the asymptotic stability of optimal filters with respect to perturbations to their initial condition for systems with ergodic limiting behavior. So now we summarize the material we need from [8], restricting to a specific situation which is appropriate to our purpose.

4.1 Stability of filters with respect to their initial condition.

Consider a filtering model specified by a signal-observation pair (Z, Y) such that:

- the signal $Z = (Z_t\,,\, t \geq 0)$ is a continuous, $\mathbb{R}^n$-valued homogeneous Markov process with some initial distribution π on $\mathbb{R}^n$,
- the $\mathbb{R}^p$-valued observation process Y is defined by

$$Y_t = \int_0^t H(Z_s)\,ds + W_t\,,\ t \geq 0\,,$$

  where W is a $\mathbb{R}^p$-valued standard Brownian motion independent of Z and H is a bounded continuous function from $\mathbb{R}^n$ into $\mathbb{R}^p$.

Let us work as before on the canonical stochastic basis, which here can be taken as $\Omega = \mathcal{C}(\mathbb{R}^+;\mathbb{R}^n) \times \mathcal{C}(\mathbb{R}^+;\mathbb{R}^p)$. We denote by $\mathbb{P}^\pi$ the distribution on Ω corresponding to the above model specified by the initial condition π. We denote also by $\mathbb{E}^\pi[\,.\,]$ an expectation computed with respect to $\mathbb{P}^\pi$ so that the optimal filter Π_t^π for Z_t corresponding to the prior distribution π is defined by

$$\Pi_t^\pi(\psi) = \mathbb{E}^\pi[\psi(Z_t)|\mathcal{Y}_t],\ \ \psi \in \mathcal{M}_b(\mathbb{R}^n).$$

Given two different initial conditions π_0 (the true prior distribution say) and $\bar{\pi}_0$ (an incorrect prior distribution), Ocone and Pardoux [8] studied the long time behavior of the difference $\Pi_t^{\pi_0}(\psi) - \Pi_t^{\bar{\pi}_0}(\psi)$ under the probability $\mathbb{P}^{\pi_0}$. They introduce some conditions which in the present case may be stated as follows. Let $\mathcal{C}_b(\mathbb{R}^n)$ be the set of those functions in $\mathcal{M}_b(\mathbb{R}^n)$ which are continuous and let $(S_t, t \geq 0)$ denote the transition semigroup of the process Z. Assume that:

- (H_1) $(S_t, t \geq 0)$ is a strongly continuous semigroup satisfying the Feller property i.e.

$$\text{for all } t \geq 0\,, \quad S_t(\mathcal{C}_b(\mathbb{R}^n)) \subset \mathcal{C}_b(\mathbb{R}^n)\,,$$

  and admitting a unique invariant measure μ such that for all $\varphi \in \mathcal{C}_b(\mathbb{R}^n)$

$$\limsup_{t \to +\infty} \int_{\mathbb{R}^n} |S_t\varphi(z) - \mu(\varphi)|\, \mu(dz) = 0.$$

Then, denoting by πS_t the distribution of the random variable Z_t under $\mathbb{P}^\pi$, it is said that $(S_t, t \geq 0)$ forgets π for μ if

- $(H_2(\pi))$ $\pi S_t \to \mu$ weakly as $t \to +\infty$.

In what follows R^π denotes the marginal distribution of $\mathbb{P}^\pi$ on $\Omega_2 = \mathcal{C}(\mathbb{R}^+; \mathbb{R}^p)$ i.e. the distribution of the observation process Y corresponding to the initial distribution π for the signal Z.

Then, applying [8, Theorem 3.2] to the just described situation, the following statement holds:

Theorem 4.1 *Assume that*

(i) $(S_t, t \geq 0)$ *satisfies* $(H1)$*,*

(ii) $(H_2(\pi_0))$ *and* $(H_2(\bar{\pi}_0))$ *are both satisfied,*

(iii) R^{π_0} *is absolutely continuous with respect to* $R^{\bar{\pi}_0}$*.*

Then, for every continuous $\psi \in \mathcal{M}_b(\mathbb{R}^n)$*,*

$$\lim_{t \to +\infty} \mathbb{E}^{\pi_0}[|\Pi_t^{\pi_0}(\varphi) - \Pi_t^{\bar{\pi}_0}(\varphi)|^2] = 0.$$

4.2 Asymptotic optimality of the approximate filter.

Of course we may look at model (56) as a signal-observation model of the type which has just been discussed in Section 4.1. The process $Z = (\zeta', X')'$ is the $\mathbb{R}^{k+d}$-valued Markovian signal with initial condition $\pi_0 = \mathbb{P}_{(\zeta_0, X_0)}$. The exact (resp. approximate) filter Π_t (resp. Π_t^0) is the optimal filter $\Pi_t^{\pi_0}$ (resp. $\Pi_t^{\bar{\pi}_0}$) corresponding to the true (resp. incorrect) prior distribution π_0 (resp. $\bar{\pi}_0 = \delta_0 \otimes \nu_0$ where ν_0 is the distribution of X_0, i.e. the second marginal of π_0).

Now we shall prove that the Theorem 4.1 can be applied to obtain the asymptotic optimality of the approximate filter Π_t^0 in the following sense:

Proposition 4.1 *Assume that in the model (56), the conditions* $(A_0) - (A_1') - (A_{2,1}^*) - (A_{2,2}^*) - (A_3^*)$ *are fulfilled. Then for every continuous* $\varphi \in \mathcal{M}_b(\mathbb{R}^d)$,

$$\lim_{t \to +\infty} \mathbb{E}[|\Pi_t(\varphi) - \Pi_t^0(\varphi)|^2] = 0. \tag{58}$$

The proof consists to show that the conditions $(i) - (iii)$ of Theorem 4.1 are all satisfied. This will be done through two lemmas. At first let us notice that, from Hasminski [3], condition (A_3^*) is actually a sufficient condition for the ergodicity of a Markov diffusion process X^0 associated with the first stochastic differential equation in (57). Under this condition the corresponding semigroup $(S_t^0, t \geq 0)$ admits a unique invariant measure μ^0 and $(S_t^0, t \geq 0)$ forgets any probability measure ν on $\mathbb{R}^d$ for μ^0 i.e.

$$\nu S_t^0 \to \mu^0 \text{ weakly as } t \to +\infty\,, \tag{59}$$

where νS_t^0 denotes the law of the random variable X_t^0 when the distribution of X_0^0 is ν.

Then the first lemma says that conditions $(i) - (ii)$ in Theorem 4.1 are satisfied.

Lemma 4.1 *Assume that conditions* $(A_0) - (A_1') - (A_{2,1}^*) - (A_{2,2}^*) - (A_3^*)$ *are satisfied. Let* $(S_t, t \geq 0)$ *be the transition semigroup of the diffusion process* $Z = (\zeta', X')'$ *generated by equations (56). Then* $(S_t, t \geq 0)$ *is a strongly continuous semigroup satisfying the Feller property and admitting a unique invariant measure* μ *defined by* $\mu = \delta_0 \otimes \mu^0$ *which satisfies* (H_1)*. Moreover, for* $\pi_0 = \mathbb{P}_{(\zeta_0, X_0)}$ *and* $\bar{\pi}_0 = \delta_0 \otimes \nu_0$*, conditions* $(H_2(\pi_0))$ *and* $(H_2(\bar{\pi}_0))$ *are fulfilled.*

Proof Under our assumptions, it is well known that $(S_t)_{t\geq 0}$ is a strongly continuous semigroup satisfying the Feller property. Moreover, from the discussion above about the consequences of (A_3^*) for the first equation in (57), it is straightforward to show that the measure $\mu = \delta_0 \otimes \mu^0$, where μ^0 is the unique invariant measure of $(S_t^0, t \geq 0)$, is an invariant measure of $(S_t, t \geq 0)$. The uniqueness of μ^0 implies the uniqueness of μ and from (59), (H_1) can be easily obtained.
On the other hand, since the law $\pi_0 S_t$ of Z_t under $\mathbb{P}^0 = \mathbb{P}^{\bar{\pi}_0}$ coincides with the law $\delta_0 \otimes (\nu_0 S_t^0)$ on $\Omega_0 \times \Omega_1$, the condition $(H_2(\bar{\pi}_0))$ is directly obtained from (59). It remains to verify that $(H_2(\pi_0))$ is fulfilled. Since $(S_t, t \geq 0)$ admits a unique invariant measure it suffices to show that the family of laws $\{\mathbb{P}_{Z_t}\,,\, t \geq 0\}$ where $\mathbb{P}_{Z_t} = \pi_0 S_t$, is uniformly tight i.e. for any $\varepsilon > 0$, there is a compact set $K \subset \mathbb{R}^n$ such that

$$\mathbb{P}([Z_t \in K]) \geq 1 - \varepsilon\,, \text{ for all } t \geq 0\,.$$

But from $(A_{2,1}^*) - (A_{2,2}^*)$, ζ converges to 0 in probability under $\mathbb{P}$. Then, denoting $\mathbb{P}_{X_t}$ the distribution of X_t under $\mathbb{P}$, it suffices to prove that the family $\{\mathbb{P}_{X_t}, t \geq 0\}$ is uniformly tight.
Let us use the notations introduced in Section 2.1. Recall that we defined there the probability $\widetilde{\mathbb{P}}$ which is locally absolutely continuous with respect to $\mathbb{P} = \mathbb{P}^{\pi_0}$ with the local Radon-Nikodym derivative L_t. Then also $\mathbb{P}$ is locally absolutely continuous with respect to $\widetilde{\mathbb{P}}$ with the local Radon-Nikodym derivative L_t^{-1}. Then for any $R > 0$, we have

$$\mathbb{P}([|X_t| > R]) = \widetilde{\mathbb{E}}[\mathbb{1}_{\{|X_t|>R\}} L_t^{-1}],$$

But under $\widetilde{\mathbb{P}}$ the process X satisfies the stochastic differential equation

$$dX_t = b(X_t)\, dt + d\widetilde{V}_t,\ t \geq 0, X_0,$$

where $\widetilde{V}$ is a Brownian motion and X_0 has still distribution ν_0 and is independent of $\widetilde{V}$. Therefore under $\widetilde{\mathbb{P}}$ the distribution of the random variable X_t coincides with $\nu_0 S_t^0$. Hence by the Cauchy-Schwarz inequality for any $R > 0$ we get

$$\mathbb{P}([|X_t| > R]) \leq \left[\widetilde{\mathbb{P}}([|X_t| > R])\right]^{\frac{1}{2}} \left[\widetilde{\mathbb{E}}(L_t^{-2})\right]^{\frac{1}{2}},$$

or equivalently

$$\mathbb{P}([|X_t| > R]) \leq \left[\nu_0 S_t^0(\{x \in \mathbb{R}^d \,:\, |x| > R\})\right]^{\frac{1}{2}} \left[\mathbb{E}(L_t^{-1})\right]^{\frac{1}{2}}. \tag{60}$$

But, taking into account the independence of ζ and (V, W) under $\mathbb{P}$, conditioning on ζ it is easy to check that

$$\mathbb{E}(L_t^{-1}) = \mathbb{E}(\exp < K >_t).$$

Then making use of assumptions $(A_0) - (A_1') - (A_{2,1}^*) - (A_{2,2}^*)$ it is readily seen that there exists some positive constant C such that

$$\mathbb{E}(L_t^{-1}) \leq C < +\infty. \tag{61}$$

Moreover, assumption (A_3^*) implies that the family of distributions $\{\nu_0 S_t^0, t \geq 0\}$ is uniformly tight. Therefore, for given $\varepsilon > 0$ and $C > 0$, there exists a $R > 0$ such that

$$\nu_0 S_t^0(\{x \in \mathbb{R}^d \,:\, |x| > R\}) \leq \frac{\varepsilon^2}{C}, \text{ for all } t \geq 0\,.$$

Hence, due to (60) and (61), for this R it follows that

$$\mathbb{P}([|X_t| > R]) \leq \varepsilon, \text{ for all } t \geq 0\,.$$

This means that the family of laws $\{\mathbb{P}_{X_t}, t \geq 0\}$ is uniformly tight and we can conclude that $(H_2(\pi_0))$ is fulfilled. □

Finally the next lemma says that the condition (iii) in Theorem 4.1 is satisfied.

Lemma 4.2 *Assume that conditions $(A_0) - (A_1') - (A_{2,1}^*) - (A_{2,2}^*)$ are satisfied. Then, for $\pi_0 = \mathbb{P}_{(\zeta_0, X_0)}$ and $\bar{\pi}_0 = \delta_0 \otimes \nu_0$, the distribution R^{π_0} is absolutely continuous with respect to $R^{\bar{\pi}_0}$.*

Proof Recall that R^{π_0} and $R^{\bar{\pi}_0}$ are the marginal distributions on Ω_2 of $\mathbb{P} = \mathbb{P}^{\pi_0}$ and $\mathbb{P}^0 = \mathbb{P}^{\bar{\pi}_0}$ respectively. Again we use the notations introduced in Section 2.1. The probability $\widetilde{\mathbb{P}}$, already used in the proof of Lemma 4.1 has the same marginal distribution on $\Omega_1 \times \Omega_2$ as the probability $\mathbb{P}^0 = \mathbb{P}^{\bar{\pi}_0}$ and hence its marginal distribution on Ω_2 is nothing but $R^{\bar{\pi}_0}$. Hence to prove the statement in the lemma it is sufficient to show that in fact $\mathbb{P}$ is absolutely continuous with respect to $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$. But since $\mathbb{P}$ is locally absolutely continuous with respect to $\widetilde{\mathbb{P}}$ with the local Radon-Nikodym derivative L_t^{-1}, thanks to [4, Proposition III.3.5], to conclude it is enough to show that the density process L_t^{-1} is a square-integrable martingale with respect to $\widetilde{\mathbb{P}}$ i.e.

$$\sup_t \widetilde{\mathbb{E}}(L_t^{-2}) < +\infty,$$

or equivalently

$$\sup_t \mathbb{E}(L_t^{-1}) < +\infty.$$

Actually this is true because of the bound (61) obtained in the proof of Lemma 4.1 and so the statement in the present lemma holds. □

Remark 4.1 It is interesting to emphasize that the constant c_0 appearing in the condition $(A_{2,2}^*)$ which ensures the existence of an exponential moment for ζ_0 depends on the convergence rate λ of the flow ϕ appearing in $(A_{2,1}^*)$ and the Lipschitz constants of b_ζ and h_ζ in (A_1'). In particular, not surprisingly, it appears that the larger is λ, the smaller is the constant c_0.

Remark 4.2 With slight modifications in the proof of Proposition 4.1, one can show that the same asymptotic result holds for any initial law of the form $\bar{\pi}_0 = \delta_0 \otimes \nu$ in place of $\bar{\pi}_0 = \delta_0 \otimes \nu_0$ provided that the incorrect prior distribution ν for X_0 dominates the true one ν_0 i.e. $\nu_0 << \nu$. So for two different initial conditions of this form the asymptotic behavior of the approximate filter is unchanged. In the particular case of the filtering model (57), this reduces to the fact that, under the assumptions "b Lipschitzian", "h continuous bounded" and (A_3^*), a filter initialized with an erroneous prior distribution ν such that $\nu_0 << \nu$ has the same asymptotic behavior as the optimal filter (initialized with ν_0). Actually that property in model (57) is an immediate consequence of [8, Theorem 3.2 and Remark 3.3].

Acknowledgments : We are very grateful to Marina Kleptsyna for valuable discussions concerning ergodic properties of diffusion processes.

References

[1] V. E. Beněs and I. Karatzas : *Estimation and control for linear, partially observable systems with non–Gaussian initial distribution*, Stochastic Process. Appl., 14, 1983, pp. 233–248.

[2] R.J. Elliott : *Stochastic Calculus and Applications*, Springer-Verlag, New-York, 1982.

[3] R. Z. Hasminski : *Stochastic stability of differential equations*, Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands and Rockville, Maryland ,USA, 1980.

[4] J. Jacod and A. N. Shiryaev : *Limit theorems for stochastic processes*, Springer–Verlag, 1987.

[5] H. Kwakernaak and R. Sivan : *Linear Optimal Control Systems*, Wiley-Interscience, New-York, 1972.

[6] A. Le Breton and M.-C. Roubaud : *Asymptotic behavior of reduced-order filters*, Rapport de Recherche INRIA, No. 3311, 1997.

[7] A. M. Makowski : *Filtering formulae for partially observed linear systems with non–Gaussian initial conditions*, Stochastics, 16, 1986, pp. 1–24.

[8] D. Ocone and E. Pardoux : *Asymptotic stability of the optimal filter with respect to its initial condition*, SIAM J. Control and Optimization, Vol. 34, No. 1, 1996, pp. 226–243,

[9] E. Pardoux : *Filtrage non linéaire et équations aux dérivées partielles stochastiques associées*, In P.L. Hennequin, editor, Ecole d'été de Probabilités de Saint–Flour XIX – 1989, Vol. 1464 of Lecture Notes in Mathematics, pp. 67–163, Springer Verlag, Berlin, 1991.

[10] M. Zakai : *On the optimal filtering of diffusions processes*, Zeit. Wahr. Verv. Geb., 11, 1969, pp. 230–243.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399